

Multidimensional Intrusion Detection System for IEC 61850 based SCADA Networks

Yi Yang, *Member, IEEE*, Lei Gao, Yu-Bo Yuan, Kieran McLaughlin, Sakir Sezer, *Member, IEEE*, Yan-Feng Gong, *Senior Member, IEEE*

Abstract--Emerging cybersecurity vulnerabilities in supervisory control and data acquisition (SCADA) systems are becoming urgent engineering issues for modern substations. This paper proposes a novel intrusion detection system (IDS) tailored for cybersecurity of IEC 61850 based substations. The proposed IDS integrates physical knowledge, protocol specifications and logical behaviors to provide a comprehensive and effective solution that is able to mitigate various cyberattacks. The proposed approach comprises access control detection, protocol whitelisting, model-based detection, and multi-parameter based detection. This SCADA-specific IDS is implemented and validated using a comprehensive and realistic cyber-physical test-bed and data from a real 500kV smart substation.

Index Terms-- Smart substation, SCADA, cybersecurity, IEC 61850, intrusion detection.

I. INTRODUCTION

IEC 61850 [1] based supervisory control and data acquisition (SCADA) systems play a significant and increasingly critical role in smart grid operation, becoming more complex and interconnected as state-of-the-art information and communication technologies (ICT) are adopted. The increased complexity and interconnection of SCADA systems have exposed them to a wide range of cybersecurity threats, which may lead to serious physical damage [2].

In recent years, malicious cybersecurity incidents have occurred in industrial control systems around the world. For instance, in July 2010 the *Stuxnet worm* that attacked Iranian nuclear facilities is the most famous malware attack to damage an industrial infrastructure directly [3]; in December 2015, a coordinated intentional cyberattack via the *BlackEnergy* malware was directly responsible for power outages for at least 80,000 customers in western Ukraine. The incident is the first known power outage caused by a cyberattack [4]. *Stuxnet* and

BlackEnergy have demonstrated that “security by obscurity” is no longer an adequate scheme for critical infrastructures. Many governments and government agencies have expressed concern at the possibility of catastrophic damage to their critical infrastructures from Stuxnet-like or BlackEnergy-like attacks in the future.

As these threats have emerged, electrical utilities have found that existing IT-specific security methodologies are not fully compatible with IEC 61850 based SCADA operation scenarios. For example, traditional IT security appliances such as firewalls and intrusion detection system (IDS) are generally unable to interpret the application layer data for such communications, either for a single packet, or at a session layer, where the state of a connection should be monitored for inconsistencies. In addition, to provide an accurate analysis of the network communications, the analyzing system needs to have some knowledge of the underlying physical infrastructure in order to process decisions about whether observed patterns of communication are benign or malicious. While generic IT communications are heterogeneous and widely varied in nature, a cyber-physical system has a certain structure and communication patterns that should be used to support detection of suspicious activities. Furthermore, although the IEC 62351 [5] standard defines a framework for the provision of cybersecurity for the IEC 61850 protocol, major manufacturers do not generally implement adequate security in their intelligent electronic devices (IEDs) [6]. With vendors slow to respond, it has become essential that utilities are able to fill this security gap to enable them to detect and mitigate again emerging threats. However, contemporary intrusion detection approaches are generally inadequate for application to this domain. Consequently, the contribution of the presented research is a network intrusion detection system tailored to respond to cyberattacks that intend to exploit and disrupt systems reliant on IEC 61850, which is likely to be the dominant protocol in emerging smart grid systems. Furthermore, this research has been informed by, developed for, and validated within the context of a real substation environment.

Much research has been proposed in intrusion and anomaly detection targeted for SCADA systems [7]-[18]. However, research on cost-effective IDS for IEC 61850 smart substations is still in an early stage of development [19]-[22]. Cheung *et al.* [8] believed that model-based monitoring to detect unknown attacks is more feasible in SCADA systems

This work was supported in part by Natural Science Foundation of Jiangsu Province under Grant Number BK20140114, and the European FP7 project SPARKS under Grant Number 608224.

Yi Yang, Lei Gao and Yu-Bo Yuan are with State Grid Jiangsu Electric Power Company Research Institute, Nanjing, 210000 China (e-mail: yyang09@qub.ac.uk; gaolei_seu@163.com; yybseu@sohu.com).

Kieran McLaughlin and Sakir Sezer are with the School of Electronics, Electrical Engineering and Computer Science, Queen’s University Belfast, Belfast, BT9 5AH U.K. (e-mail: kieran.mclaughlin@ee.qub.ac.uk; s.sezer@ecit.qub.ac.uk).

Yan-Feng Gong is with School of Electrical Engineering, North China Electric Power University, Beijing, China 102206 (e-mail: yanfeng.gong@ncepu.edu.cn).

than in general IT networks, using protocol-level modes, communication-pattern-based detection and a learning-based approach. Unfortunately, no quantitative results were obtained from this work nor detailed analysis regarding experimental validation. Carcano *et al.* [9] proposed critical state-based IDS for SCADA based on the Modbus protocol in a power plant. However, this system can only detect a limited class of attacks against programmable logic controller (PLC) systems. Fovino *et al.* [10] also utilized critical states in IDS supporting Modbus and DNP3. Barbosa *et al.* [11] adopted a network flow whitelisting based intrusion detection approach for the security of SCADA systems. The flow whitelist in the proposed approach is learned by capturing network traffic at two water treatment plants and at an electric-gas utility. However, this detection approach did not consider protocol specifications and features such as IEC 61850. Premaratne *et al.* [12] used a rule-based IDS for an IED based on IEC 61850 in Snort parlance. The Snort rules are obtained from experimental data based upon simulated cyberattacks without considering the protocol's specification. The proposed blacklist approach is shown to detect known attacks effectively. However, blacklists are typically not effective against unknown threats or undiscovered vulnerabilities, also called zero-day attacks. Kwon *et al.* [14] proposed a behavior-based IDS to detect anomalous events by statistical analysis of IEC 61850 based substation network traffic, limited to manufacturing message specification (MMS) and generic object oriented substation event (GOOSE) messages. However, most statistical intrusion methods generate false negatives which miss real attacks. Hong *et al.* [15] presented a host- and network-based anomaly detection system to detect simulated attacks in substations. However, this anomaly detection is limited to the multicast protocols, i.e., GOOSE and sampled measure value (SMV). Yoo *et al.* [18] proposed an anomaly-detection system for the IEC 61850 protocols (MMS and GOOSE) including pre-processing, normal-behavior learning and anomaly detection. However, its detection accuracy still needs to be improved in order to apply it in the real substation. Much more in-depth insight into integrating physical knowledge, protocol specifications and logical behaviors with SCADA-specific IDPS is urgently required for cybersecurity of IEC 61850 based control systems.

In response to the challenge represented by cyber vulnerabilities in IEC 61850 smart substations [23], this paper proposes a novel IDS. The comprehensive SCADA-specific IDS is tailored for cybersecurity of IEC 61850 based SCADA networks. It consists of access control detection, protocol whitelisting detection, model-based detection, and multi-parameter based detection. This final component, based on multiple parameters, utilizes inspection of communications at the application layer in order provide exceptionally fine grained monitoring of system commands for anomalies. This SCADA-specific IDS is implemented and validated using a realistic

cyber-physical test-bed of a 500kV smart substation.

Section II presents the technical background. Section III proposes a novel intrusion detection system for IEC 61850 based SCADA networks. Section IV discusses the implementation approach of the proposed SCADA-IDS. In Section V, a SCADA-specific cybersecurity test-bed is presented to investigate potential intrusions, exemplify and validate the proposed SCADA-IDS.

II. BACKGROUND

This section contains the brief introduction of IEC 61850 and substation configuration description language from the viewpoint of supporting the proposed IDS for IEC 61850 based SCADA networks.

A. IEC 61850

The abstract data models defined in IEC 61850 can be mapped to many protocols. Current mappings in this standard are mainly to MMS, GOOSE, and SMV [24], [25]. The MMS protocol is applied in the station level based on the client/server model, which runs over TCP/IP networks. The GOOSE and the SMV protocols are both based on publish/subscription mechanism in the substation local area network (LAN) using high speed switched Ethernet [31]. The IEC 61850 protocol stack is shown in Fig. 1.

In terms of the transport layer of the MMS protocol stack in Fig. 1, international standards organization (ISO) transport (ISO/IEC 8073) means connection oriented transport protocol (COTP), and RFC 1006 stands for ISO transport services on top of the TCP (TPKT) (the TCP port for TPKT traffic is 102). In Fig. 1, ACSI, abstract communication service interface, defines the virtual interface to an IED providing abstract communication services [1].

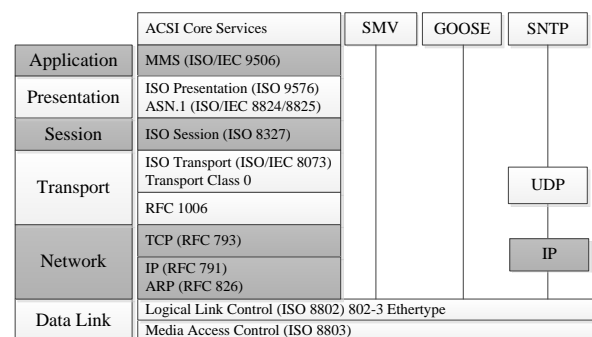


Fig. 1. IEC 61850 protocol stack

The GOOSE/SMV datagrams comply with ISO/IEC 8802-3 in the data link layer. The ISO/IEC 8802-3 frame format for GOOSE/SMV packets is illustrated in Fig. 2. The *destination address* has six octets corresponding to an Ethernet MAC multicast address. The *source address* is a unicast MAC address. According to IEEE 802.1Q, the *priority/VLAN tag* is used to separate high priority and time critical traffic for

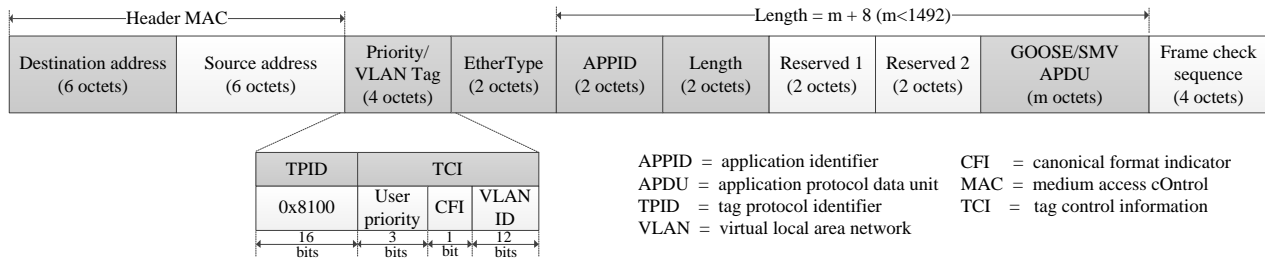


Fig. 2. ISO/IEC 8802-3 frame format for GOOSE/SMV datagrams

protection relevant applications from low priority traffic. Abstract syntax notation one (ASN.1) in relation with basic encoding rules (BER) is used for encoding and decoding of the GOOSE/SMV messages for transmission on ISO/IEC 8802-3, which has the format of a triplet TLV (Tag, Length, Value) [24], [25]. The *destination address*, *user priority*, *VLAN ID*, *application identifier* (APPID) and several fields of *application protocol data unit* (APDU) are configured in a substation configuration description (SCD) file for a practical smart substation.

B. Substation Configuration Description Language (SCL)

SCL files are used to exchange the configuration data, such as IED capability description (ICD) and SCD [26] [27]. The SCL is able to describe a substation configuration and all IEDs configurations in the substation using object models. It also specifies a unified and standardized format for configuring the substation and related IEDs. Therefore, the security, reliability and interoperation of smart substations are based on SCL configuration files. SCL, based on XML 1.0, defines specific syntax structures using XML Schema. A typical SCL file contains five elements, such as *Header*, *Substation*, *Communication*, *IED*, and *DataTypesTemplates*. An SCL file is a typical tree structure, as shown in Fig. 3. An SCL element is a node of the tree, and nested elements are child nodes of the tree. In Fig. 3, LN, DOType, DAType represent the logic node, type of data object, and type of data attribute, respectively.

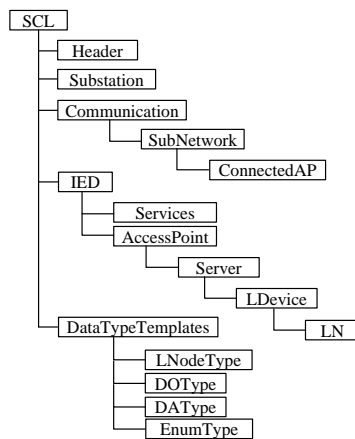


Fig. 3. Structure of SCL configuration file

III. PROPOSED IDS FOR IEC 61850 BASED SCADA

Based on recent related cyberattacks such as Havex, Stuxnet and Ukraine, the motivation for the proposed network IDS is

to detect SCADA-specific behaviors carried out by an intruder who has already gained a foothold in the network due to an infected human machine interface (HMI), engineering laptop, or a similar initial vector. These initial infections typically exploit IT software vulnerabilities unrelated to the core control system. However at this point the intruder is likely to attempt reconnaissance activity on the SCADA network, to scan the network, enumerate hosts and devices, and gather intelligence about devices of interest, for example IEDs, etc. Unless they have well established intelligence from some other source they may well attempt some fuzzing activity on the network to establish responses from devices of interest. Assuming that cybersecurity preventative measures have failed and allowed this intrusion, it is now vital that a further layer of detection can react to the abovementioned intrusion activities in the highly sensitive SCADA network.

A SCADA-IDS for IEC 61850 smart substations is therefore proposed as an effective tool to identify both external malicious attacks and internal unintended misuse. This novel mechanism blends physical knowledge and behavioral logic of power systems with emerging IT security approaches. The proposed IDS approach consists of four dimensions: 1) access-control detection; 2) protocol whitelisting detection; 3) model-based detection; 4) multi-parameter based detection.

The theoretical basis supporting this scheme is grounded in the well-established cyber security principle of defense in depth [30]. This theory first recommends establishing a network perimeter policed by standard security controls such as firewalls. For the scenario investigated in this paper a logical network perimeter can be formed around the digital substation, which contains a secure zone of IEC 61850 related communications. The next stage to ensure defense in depth is to establish monitoring mechanisms within the secure zone that can detect breaches and failures of security controls, e.g. an attacker penetrating a misconfigured firewall, or bypassing the firewall completely by launching an attack from a malware-infected laptop that has been directly connected to the substation LAN by an engineer. Once an intruder has established a presence in the target substation network, automated or manual activities will be initiated, ranging from basic network scans, to “fuzzing” and deliberately crafting packets to attempt to gain a response from an IED, or to cause a specific command to be executed. Each of these actions will not be prevented by perimeter firewalls, because they occur inside the perimeter of the secure zone. Therefore the proposed IDS IEC 61850 has been designed to detect the multiple layers of activity and complexity of communications that may be

generated by a successful intruder. Four methods to support this multidimensional defense in depth IDS approach have been implemented, and will now be explained.

A. Access-Control Detection (ACD)

The ACD is a kind of access-control whitelist strategy including medium-access control (MAC) addresses in the Ethernet layer, IP addresses in the network layer and ports in the transport layer. The TCP port for IEC 61850 traffic is <102>. If any of the addresses or ports is not in the corresponding whitelist, the detector will take pre-configured action, e.g., alert in the IDS mode, block in IPS (intrusion prevention system) mode, and log the detection results. That is,

$$AC \notin \{AC_{wl}\} \rightarrow Actions(alert / block, log) \quad (1)$$

where $AC = MAC_{src}, MAC_{dst}, IP_{src}, IP_{dst}, Port_{src}, Port_{dst}$ and AC_{wl} represents corresponding whitelist set. $MAC_{src}, MAC_{dst}, IP_{src}, IP_{dst}, Port_{src}, Port_{dst}$ mean source and destination MAC, source and destination IP, as well as source and destination ports, respectively.

Each host or device in a SCADA system has a unique <IP, MAC> match. If the device has not been replaced with new hardware and the same IP address of the device is detected from two or more MAC addresses, it means that a spoofing attack may be happening. Malware attempting to communicate out to a command and control server can also generate unexpected address and network activities, particularly after the initial infection stage.

B. Protocol Whitelisting Detection (PWD)

The protocol whitelisting detection refers to layers 2-7 in terms of the open systems interconnection (OSI) model, and deals with various protocols of smart substation networks, such as MMS, COTP, TPKT, simple network time protocol (SNTP), GOOSE, SMV, and IEEE 1588. A typical substation based on IEC 61850 consists of a station bus and a process bus. In terms of the station bus, the detector can be set to allow communication traffic complying with MMS/COTP/TPKT/SNTP. In terms of the process bus, the detector will only allow GOOSE/SV/IEEE 1588 traffic. In different scenarios, the detector can be set to support specific protocols. For example, when the IDS is deployed at the process bus of the smart substation, this detector only allows GOOSE/SV/IEEE 1588 traffic, otherwise, it will generate an alert message for the suspicious traffic.

C. Model-based Detection (MBD)

The proposed model-based detection approach analyses SCD files and normal IEC 61850 traffic contents, defines normal and correct behavior models using in-depth protocol analysis, and compares profiles of benign behaviors against observed traffic to identify anomalous deviations. A model-based anomalous behavior detection approach has the potential to detect as-yet unknown attacks. Compared with traditional IT networks, SCADA networks in smart substations have distinguishing characteristics such as regular traffic flows and predictable behavior patterns, which potentially simplifies the

specification of behavior models. The proposed MBD has the potential to identify malicious attacks or unintended anomalies both in the station bus and the process bus.

1) MBD for Station Bus

In the station bus, the anomalous behavior detection is based on ACSI (mapping to MMS) or SNTP. The detection models are defined as follows,

a) Report Service Model

In the SCD file, the maximum number of instantiable report control blocks of each IED has been configured. The proposed report service model defines the maximum number of instantiable report control blocks for each IED as a detection rule. If the MBD identifies abnormal connection requests that could occupy all the instantiable report control blocks of the IED, it will alert a suspicious denial-of-service (DoS) attack and log the detection results.

b) Association Service Model

The proposed association service model defines the maximum number of IEC 61850 clients that can be connected. If the MBD detects abnormal connection requests to the clients, it will generate an alert and log the detection result.

c) Setting Service Model

The proposed setting service model defines that only an IEC 61850 client is allowed to modify a setting. If this model is violated, the MBD will generate actions (alert and log).

d) File Transfer Model

The ACSI *GetFile* service is used by a client to transfer the contents of a file from the server to the client. The ACSI *GetFileAttributeValues* service is used by a client to obtain the name and attributes of a specific file in the server's file store [24]. The proposed file transfer model defines an IEC 61850 client can only transfer a single file. If this rule is violated, it will generate an alert and log the detection result.

e) SNTP Model

In the substation network, SNTP [28] is used to accomplish time synchronization via LAN communication. The SNTP traffic adopts the user datagram protocol (UDP) in the transport layer. In terms of the SNTP traffic, the port number of UDP connection to an IEC 61850 server should be <123>. If the port number of the SNTP traffic is not <123>, the MBD will trigger an alarm and save the result in the log file.

f) Time-Related Model

Critical control commands have time-related constraints such as time interval limit and frequency limit. If the same legitimate command is sent too frequently, it may violate the following rules. In each case the detector will initiate some actions (alert and log).

$$CV(n) - CV(n-1) < T \rightarrow Actions(alert, log) \quad (2)$$

where CV is a control command, n is a positive integer ($n > 1$), and T is the limit of time interval.

$$\frac{CV(n) - CV(1)}{n-1} > F \rightarrow Actions(alert, log) \quad (3)$$

where F represents the frequency limit.

2) MBD for Process Bus

In the process bus, the model-based detection is based on GOOSE and SMV protocol specifications. The GOOSE APDU has twelve fields such as *gocbRef* (control block reference), *timeAllowedToLive*, *datSet* (data set reference), *goID* (GOOSE ID), *t* (event timestamp), *StNum* (state number), *SqNum* (sequence number), *test* (test identifier), *confRev* (configuration revision), *ndsCom* (needs commissioning), *numDatSetEntries* (number of data set entries) and *allData* [24]. According to IEC 61850-9-2, the SMV datagram adopts ISO/IEC 8802-3 in the data link layer, similar to the GOOSE datagram. The SV APDU has five fields such as *svID* (SMV control block ID), *smpCnt* (sample counter), *confRev* (configuration revision), *smpSynch* (sample synchronization), and *seqData* (sequence of data). The part of proposed detection models are defined as follows,

a) Destination Address Model

The destination ISO/IEC 8802-3 multicast address is configured for the transmission of GOOSE/SMV in the SCD file (<Communication>→<SubNetwork>→<ConnectedAP>). The destination address fields (6 octets) of a GOOSE packet and a SMV packet start with four octets (01-0C-CD-01) and (01-0C-CD-04), respectively. The destination address models for GOOSE and SMV are shown in (4) and (5), i.e.

$$\forall P \in P_{GOOSE} \Rightarrow DstAdrField(P) \in [01-0C-CD-01-00-00, 01-0C-CD-01-01-FF] \quad (4)$$

where P is the captured packet in the process bus, P_{GOOSE} represents GOOSE packets and $DstAdrField$ represents the value of the destination address field in the ISO/IEC 8802-3 frame format.

$$\forall P \in P_{SMV} \Rightarrow DstAdrField(P) \in [01-0C-CD-04-00-00, 01-0C-CD-04-01-FF] \quad (5)$$

where P_{SMV} represents SMV packets.

b) TPID Field Model

The tag protocol identifier (TPID) field (2 octets) shows the Ethertype assigned for 802.1Q Ethernet encoded frames. The value of the *TPID* field in the GOOSE/SMV packet shall be 0x8100, i.e.

$$\forall P \in P_{GOOSE/SMV} \Rightarrow TPIDField(P) = 0x8100 \quad (6)$$

where *TPIDField* means the value of the *TPID* field, and $P_{GOOSE/SMV}$ represents GOOSE or SMV packets.

c) EtherType Field Model

The *EtherType* field (2 octets) of ISO/IEC 8802-3 is registered by the IEEE authority. The assigned *EtherType* values for GOOSE and SMV are 0x88B8 and 0x88BA, respectively, i.e.

$$\forall P \in P_{GOOSE} \Rightarrow EthTField(P) = 0x81B8 \quad (7)$$

where *EthTField* is the value of the *EtherType* field.

$$\forall P \in P_{SMV} \Rightarrow EthTField(P) = 0x81BA \quad (8)$$

d) Priority Field Model

The priority field (3 bits) model defines the priority values of GOOSE and SMV packets. The default value for

GOOSE/SMV is 4, which is also configured in the SCD file. The *priority* value should be from 0 to 7, i.e.

$$\forall P \in P_{GOOSE/SMV} \Rightarrow PrioField(P) \in [0, 7] \quad (9)$$

where *PrioField* is the value of the *user priority* field.

e) APPID Field Model

Each GOOSE/SMV control block has a unique *APPID* in the SCD file. The *APPID* field (2 octets) of a GOOSE packet should be 4-bit *hexadecimal*, i.e., [0000-3FFF], and that of an SMV packet should be [4000-7FFF]. This detection models are as follows,

$$\forall P \in P_{GOOSE} \Rightarrow APPIDField(P) \in [0000, 3FFF] \quad (10)$$

$$\forall P \in P_{SMV} \Rightarrow APPIDField(P) \in [4000, 7FFF] \quad (11)$$

f) Length Model

The length field (2 octets) of a GOOSE/SMV packet specifies the total number of bytes in the frame starting from *APPID* to APDU, which is equal to 8+m (m is the length of APDU, $m < 1492$). The length field model is as follows,

$$\forall P \in P_{GOOSE/SMV} \Rightarrow LengField(P) \in [8, 1500] \quad (12)$$

where *LengField* is the value of the length field.

The length of the *goID* field in the GOOSE APDU is less than 65 bytes, i.e.

$$\forall P \in P_{GOOSE} \Rightarrow LenGoIDField(P) \leq 65 \quad (13)$$

where *LenGoIDField* is the length of the *goID* field.

g) TimeAllowedToLive Field Model

The *timeAllowedToLive* field in the GOOSE APDU should be double *MaxTime* ($2T_0$). The “*MaxTime*” is typically configured as <5000> in the SCD file (<Communication>→<SubNetwork>→<ConnectedAP>→<GOOSE>→<MaxTime>). If there is no any GOOSE packet within 10000ms, this detection model will send communication interrupt alarm.

h) Tag Field Model

In the GOOSE tag field model, the *tag* values of *gocbRef*, *timeAllowedToLive*, *datSet*, *goID*, *t*, *StNum*, *SqNum*, *test*, *confRev*, *ndsCom* and *numDatSetEntries* fields of a GOOSE packet are 0x80, 0x81, 0x82, 0x83, 0x84, 0x85, 0x86, 0x87, 0x88, 0x89, and 0x8a, respectively. In the SMV tag field model, the *tag* values of *svID*, *smpCnt*, *confRev* and *smpSynch* fields of a SMV packet are 0x80, 0x82, 0x83 and 0x85, respectively.

i) SmpCnt field Models

The *smpCnt* field model specifies the values of a counter, which is incremented each time a new sample of the analogue value is taken. When the sample rate is 4000Hz (80 samples/cycle) for merging units (MUs), the values of *smpCnt* should be kept in the right order within the scope of [0, 3999], i.e.

$$\forall P \in P_{SMV} \Rightarrow SmpCField(P) \in [0, 3999] \quad (14)$$

where *SmpCField* is the value of the *smpCnt* field.

j) Correlation Models

According to the practical SCD configuration of the smart substation, the *APPID* field equals to the last two octets of the

destination address field. It can be defined as a correlation field model, i.e.

$$\begin{aligned} \forall P \in P_{GOOSE/SMV} \cdot DstAField(P)_{5,6} &= \{abcd\} \\ \in [0000, 01FF] &\Rightarrow APPIDField(P) = \{abcd\} \end{aligned} \quad (15)$$

where $DstAField(P)_{5,6}$ represents last two octets of the destination address field.

The type of the *gocbRef* field in the GOOSE APDU is visible-string comprising logical device (LD) name, logical node (LN) name, functional constraint (FC) and control block (CB) name, i.e., LD/LN\$FC\$CB. The *datSet* field in the GOOSE APDU consists of LD name, LN name and data set (DS) name, i.e., LD/LN\$DS. The default value of the *goID* field in the GOOSE APDU is similar to that of the *gocbRef* field, i.e., LD/LN\$CB. The LD/LN value in the *gocoRef* field matches with that in the *datSet* field. The control block name in the *gocoRef* field matches with that in the *goID* field. For instance, *gocbRef*: PM5001APIGO/LLN0\$GOS\$gocb1, *datSet*: PM5001APIGO/LLN0\$dsGOOSE1, *goID*: PM5001APIGO/LLN0.gocb1. The corresponding correlation filed model is presented as follow,

$$\begin{aligned} \forall P \in P_{GOOSE} \cdot APDU \in GocbField(P) &= \{LD/LNFCCB\} \\ \Rightarrow DatSField(P) &= \{LD/LN$DS\} \\ \Rightarrow GoIDField(P) &= \{LD/LN$CB\} \end{aligned} \quad (16)$$

where *GocbField*, *DatSField*, and *GoIDfield* represent the *gocbRef*, *datSet*, and *goID* fields, respectively.

The changes of state number (*StNum*) and the sequence number (*SqNum*) in the GOOSE APDU strictly comply with associated behavior patterns. The value of *StNum* shall increment when a value of *datSet* has changed in a sent GOOSE message, which shall cause the value of *SqNum* to be set to zero. When the value of *StNum* has no change, the value of *SqNum* will increment for each GOOSE transmission, but it shall roll over to 0 at its maximal value ($SqNum_{max} = 4,294,967,295$).

$$\begin{aligned} \text{If } [StNum(GP_i) = StNum(GP_{i-1})] \\ \Rightarrow SqNum(GP_i) = [SqNum(GP_{i-1}) + 1] \leq SqNum_{max} \\ \text{If } [StNum(GP_i) > StNum(GP_{i-1})] \Rightarrow SqNum(GP_i) = 0 \end{aligned} \quad (17)$$

where $StNum(GP_i)$ and $SqNum(GP_i)$ mean the *StNum* and the *SqNum* values of the i^{th} GOOSE packet, respectively.

k) Traffic based Model

According to captured traffic from practical substation scenarios, the traffic based model defines the upper and lower threshold values of the packet transfer rate per second (PPS), transfer byte size per second (BPS), the length of packets (LoP), and size of packets (SoP) as the normal traffic behaviors. This traffic detection model is as follows,

$$\begin{aligned} PPS \in [PPS_{min}, PPS_{max}] \quad BPS \in [BPS_{min}, BPS_{max}] \\ LoP \in [LoP_{min}, LoP_{max}] \quad SoP \in [SoP_{min}, SoP_{max}] \end{aligned} \quad (18)$$

where PPS_{min} and PPS_{max} represent the lower and upper threshold values of the PPS.

Any occurrences outside these proposed models are considered anomalous and suspicious. If any of the aforementioned models is violated, the MBD will generate an alert and log the detection result.

D. Multi-Parameter based Detection (MPD)

The core idea of the multi-parameter based detection is to identify possible threats against SCADA resulting from internal unintended misuse or external malicious attacks by monitoring the most operationally sensitive parameters of a smart substation. These multidimensional parameters are related to the secure and stable operation of the smart substation, such as remote measurement and remote signaling data from the station bus and the process bus in IEC 61850 substations. Multi-parameter detection strategies, such as critical switching signal correlation and key analog signal comparison, are proposed here from physical knowledge and operational experience of smart substations.

1) Remote Signaling Comparison Detector

In the IEC 61850 smart substation, intelligent terminals in the process bus apply GOOSE messages to send remote signaling data to IEDs in the bay level, and receive trip/close instructions from relay devices or monitoring and control devices. The proposed remote signaling comparison detector identifies abnormal events by comparing the GOOSE messages and associated MMS messages. For example, if a switch-in signal of a relay IED (GOOSE message) in the process level and the associated signal report (MMS message) from the station level are inconsistent, an abnormal alarm will occur, as shown in Fig. 4.

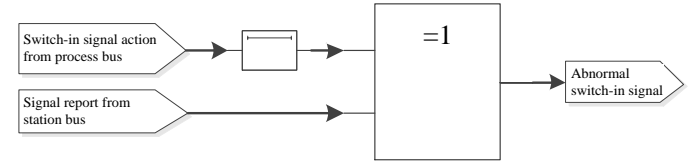


Fig. 4. An example of remote signaling comparison detector

2) Remote Measurement Comparison Detector

In IEC 61850 smart substations, merging units (MUs) have sample value models, and send SV messages to relay devices, monitoring and control devices. The remote measurement comparison detector contains two categories:

a) Range Detector

Normally, sampled measure values belong to an operational range with upper and lower boundary values, such as current (I) and voltage (U). If the measured value is outside the expected range, some actions will execute automatically, i.e.,

$$\begin{aligned} SMV(i) \notin [SMV(i)_{min} - e(i), SMV(i)_{max} + e(i)] \\ \rightarrow Actions(alert, log)(i = I, U, \dots) \end{aligned} \quad (19)$$

where $SMV(i)$ ($i = I, U, \dots$) represents different sample measure values, such as current and voltage; $[SMV(i)_{min} - e(i)$, and $SMV(i)_{max} + e(i)]$ stand for the range between the upper and lower boundary and $e(i)$ measures the tolerance.

In normal operation scenarios, the upper and lower boundaries are configured according to design and operation specifications of substations. For example, the upper and lower bus voltage boundaries for a 500(330) kV substation are set as 90% and 110% voltage rating, respectively. From the point of view of the SCADA security operation, as long as the

measured value is outside the expected range, this suspicious phenomenon should be noticed and addressed by operators in substations. Therefore, this proposed range detector may identify the abnormal incidents result from measurement errors or malicious attacks.

b) *Consistency Detector*

In practical scenarios, related double configured IEDs (A and B sets) in the bay level receive the same sampled values of MUs from associated current transformer / voltage transformer (CT/VT). The proposed consistency detector is used to detect inconsistency among configured SMV parameters of MUs and the associated MMS of multiple relay devices, such as the line relay A/B, bus relay A/B, and transformer relay A/B. Parameters for remote measurement comparison consist of the voltage, current and differential current. If the consistency detector is violated, an abnormal alarm will occur, as illustrated in Fig. 5.

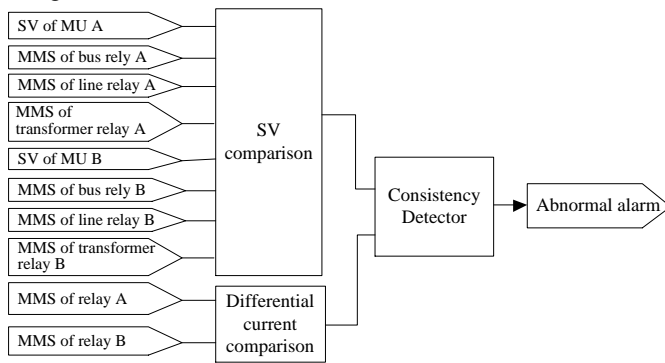


Fig. 5. Consistency detector

IV. IMPLEMENTATION

The proposed SCADA-IDS was implemented based on a *Linux* system (Ubuntu 12.04). It was deployed between the station bus and the process bus to monitor and detect SCADA traffic in both networks of the IEC 61850 based substation. It consists of five modules such as IDS configuration module, network traffic capture module, IDS process core, IDS rule module, and IDS result module, as illustrated in Fig. 6.

1) *IDS Configuration Module*: In this module, the SCADA-IDS configuration file was generated for the proposed detection approaches and rules, which includes automatic configuration information from the SCD file using the SCD parser, as well as pre-configuration information. The pre-configuration information was obtained in two ways:

a) *Self-learning from the real-time or captured normal traffic*. For example, according to the normal traffic, the authorized IP addresses, MAC addresses and port numbers were automatically learned by the IDS configuration module and added to the whitelists of access-control detection and SNTP model; based on captured MMS traffic from the station layer, the IDS configuration module has trained using 9,107,644 packets for several traffic indices, such as the upper and lower threshold values of the packet transfer rate per second (PPS) and the length of packets. The minimum value and the maximum value of packet length are 60 and 300 bytes,

respectively. The scope of PPS is [50, 400], and the real traffic of the IEC 61850 station network is demonstrated in Fig. 6.

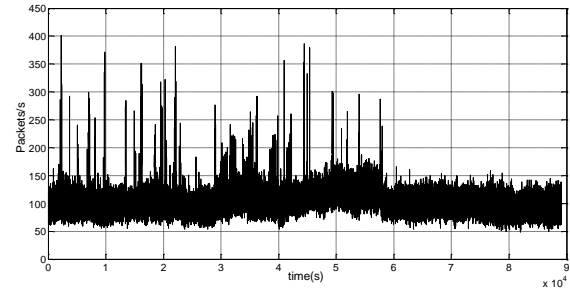


Fig. 6. Real traffic of the IEC 61850 station level network.

b) *Professional knowledge on the protocol specifications and practical operational experience*. The following are examples of the pre-configuration information.

- The implicit configuration information can be obtained from the IEC 61850 protocol and technical specifications for project implementation, with which real smart substation projects comply, e.g. in the priority field model, the GOOSE packets have priority over SMV packets in the process bus of the real smart substation, and the priority values of GOOSE and SMV packets are 6 and 4, respectively; in the destination address model, the destination address fields of a GOOSE packet and a SMV packet are set as starting with four octets (01-0C-CD-01) and (01-0C-CD-04), respectively; in the APPID field model, the *APPID* fields of GOOSE packets and SMV packets are configured with ranges of [0000-3FFF] and [4000-7FFF], respectively; in correlation model, the *APPID* field should be the last two octets of the destination address field.

- According to practical operational experience, critical control commands have time-related constraints. As an example of an interrogation command, a client in a control center might send a remote control command to request information from servers, and normally the time interval is 15 minutes.

- In normal operation of the smart substation, the SMV parameter channels of MUs and the associated MMS of multiple relay devices (A and B sets) were configured for the consistency detector. The deviation threshold of any two SMV parameters was set as 1% reference value. The threshold value of any differential current was set as 10% rating value, and deviation threshold of two differential current values was set as 5% rating value.

The above pre-configuration information is provided as a set of examples, and in practice it can be extended with much more configuration-specific data that will not be published here due to the potentially sensitive nature of some of the data.

2) *Network Traffic Capture Module*: In the module, the IEC 61850 protocol parser was developed for real-time capturing and parsing of MMS/SNTP traffic from the station bus and GOOSE/SMV traffic from the process bus. The captured actual packet capture (PCAP) files were also parsed by this module.

3) *IDS Process Core*: The IDS process core is developed based on the internet traffic and content analysis (ITACA) tool [29], which is a software platform for traffic sniffing and real-time network analysis [2]. The SCADA-specific IDS is developed in C/C++ using the ITACA platform.

4) *IDS Rule Module*: This module is the most critical component of the proposed IDS, and is developed to implement the ACD, PWP, MBD, and MPD discussed in Section III. A database is set up for the SCADA-IDS which stores critical status parameters of the SCADA system in order to realize multiple packets (cross-packet) inspection.

5) *IDS Result Module*: The detection results are demonstrated on the SCADA-IDS graphical user interface (GUI) in the substation control room and recorded in the SCADA-IDS log file, as shown in Fig. 7.

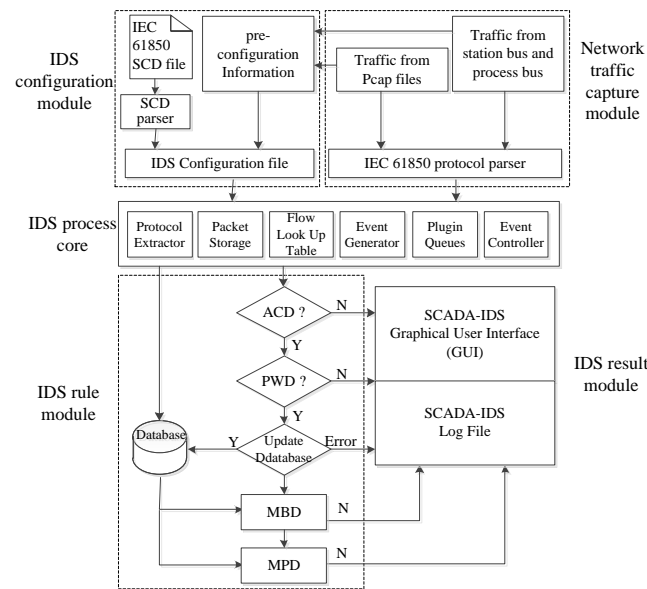


Fig. 7. Implementation of proposed IDS for smart substations.

V. TEST-BED AND EXPERIMENTAL RESULTS

A. Cyber-Physical Test-Bed

In order to investigate potential cybersecurity vulnerabilities in IEC 61850 based smart substations and verify the proposed IDS, a cyber-physical test-bed has been built in the State Grid Key Laboratory of Substation Intelligent Equipment Testing Technology in China [23], as demonstrated in Fig. 8. In this test-bed, practical IEDs, switches and monitoring system were used to replicate the SCADA network of a typical 500kV smart substation. The developed Linux-based IDS device was connected to the central station level switch and the process level switches using port mirroring.

According to the experiments in this test-bed, an infected maintenance engineer’s laptop or removable USB drive could propagate malware and launch a cyberattack tailored for smart substations, as shown by the yellow triangle in Fig. 8. The aim is to mimic the kind of attack that affected electrical utility customers in Ukraine in 2015.

B. Experimental results

In order to verify the proposed IDS approaches in this paper, an “attacker” laptop was directly connected to the station bus and the process bus to launch a number of cyberattacks in this test-bed, such as malformed packet attack, DoS attack, address resolution protocol (ARP) spoofing attack, and man-in-the-middle (MITM) attack, as depicted in the authors’ previous work [23]. The attack could equally have originated from a malware infected host on the network. In this test-bed, total 32 types of attack scenarios were exemplified in the experiment.

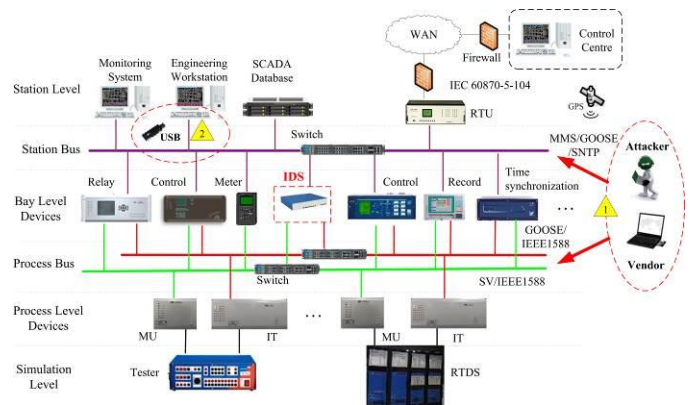


Fig. 8. Cyber-physical test-bed of IEC 61850 based smart substation

In addition, to prove the robustness of the proposed SCADA-IDS in a real environment, we conducted experiments using SCADA traffic captured from an actual operating 500kV substation based on IEC 61850. First, the normal SCADA traffic was collected as a dataset, which contains station bus traffic (2,001,928 packets), GOOSE traffic (1,757,910 packets), and SMV traffic (21,660,000 packets). Second, the real packets were retrieved using *Wireshark* and the payload data were modified using a packet revision program. 318 types of abnormal packets were generated by modifying the captured data or by injecting new malicious packets into the pre-captured PCAP file. In a real attack there are a number of ways this could be achieved, but the detectable outcome will be similar. Third, the captured traffic with abnormal packets was retransferred to the substation network. In this experiment, 32 types of proposed detection rules were integrated into the IDS rule module in Fig. 7. The effectiveness of the implemented IDS was validated with all the malicious attacks detected in the given experiment. The experimental results were recorded in a log file, and the message format in the log file is defined referring to RFC 3164. The detailed message format [28] is as follows:

```
<SEVERITY>   TIMESTAMP   DEVICE_NAME   DEVICE_TYPE
ALERT_TYPE   EVENT_DESCRIPTION   SRC_IP/SRC_MAC
(SRC_PORT)   DST_IP/DST_MAC   (DST_PORT)
```

In this case, SEVERITY represents alert severity which is described by a numerical code, e.g., 0, 1, 2 and 3 stand for EMERGENCY, ERROR, WARNING, and NOTICE, respectively. The TIMESTAMP field is the local time and is in the format of “YYYY-MM-DD HH:MM:SS.” DEVICE_NAME means the name

of a specific security device. `DEVICE_TYPE` is the type of the security device, for example, IDS. `ALERT_TYPE` represents an alert event type which is violated such as ACD, PWD, MBD, or MPD. `EVENT_DESCRIPTION` describes the detailed information of the specific security event. `SRC_MAC`, `SRC_IP`, `SRC_PORT`, `DST_MAC`, `DST_IP`, and `DST_PORT` are source MAC address, source IP address, source port, destination MAC address, destination IP address, and destination port, respectively. In terms of GOOSE/SMV detection results, only `SRC_MAC` and `DST_MAC` are required.

The logged messages generated as an output from this experiment can be understood as follows. Fig. 9 shows an alert that the suspicious state number or sequence number of the GOOSE packet is detected when a GOOSE packet is sent from the intelligent terminal A of circuit breaker 5072 (IB5072A) to the line protection A (PL5071A). According to the experiments in the cyber-physical test-bed, the false state number or sequence number of GOOSE packets in the process bus may cause rejection of relay protection. In this specific example, the real GOOSE packets were retrieved using *Wireshark* and the payload data of actual GOOSE packet was modified using a packet revision program. The associated abnormal packet is illustrated in Fig. 10. The normal `StNum` (0x0530) of the payload in Packet 448323 has become the byte (0x0531). However, the value of `SqNum` was still 0x156172, rather than zero. In the alert resulting from correlation model detection, one of MBDs is violated (discussed in Section III-C-2)).

```
<0> 2015-12-14 12:41:15 SCADA-61850-IDS IDS MBD-2-j
suspicious state number or sequence number of the
GOOSE packet **:**:00:00:10:3c **:**:cd:01:10:3c
```

Fig. 9. The MBD-2-j alert message in the log file

Fig. 11 illustrates other part of the alert messages generated due to the proposed IDS violation (described in Section III). For example, ACD-1, MBD-1-a, MBD-2-a, MPD-1, and MPD-2 specifically refer to the access-control detector, report service model, destination address model, correlation model, remote signaling comparison detector, and consistency detector, respectively. The results show how this proposed

approach can be effective against cyberattacks, since the physical effects are also detected, rather than the IT causes alone.

The indirect and valid comparisons are made between the proposed IDS and the most relevant state-of-the-art proposals, as shown in Table I. The advantages of the proposed IDS are better protocol compatibility, process time and detection accuracy.

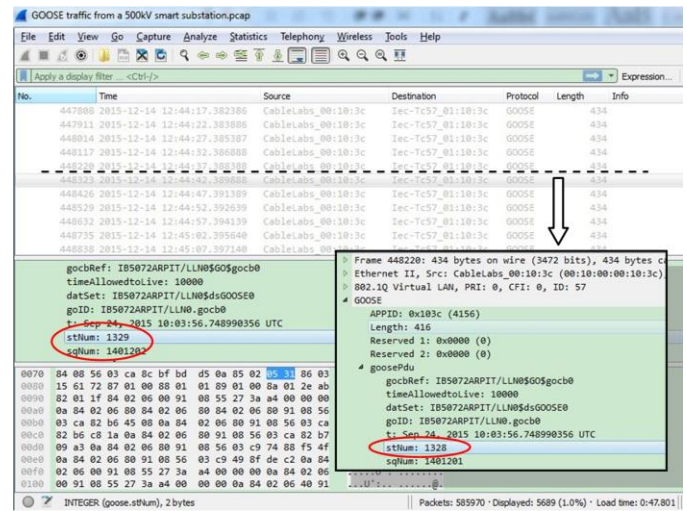


Fig. 10. Abnormal GOOSE packet detected by MBD

```
<0> 2015-12-14 18:58:23 SCADA-61850-IDS IDS ACD-1
Unauthorized Connection Attempt to a non-IEC61850
Port of a Server *.18.50.18 2218 *.18.50.215 66

<0> 2015-12-14 19:14:58 SCADA-61850-IDPS IDS MBD-1-a
Suspicious DoS attack: abnormal connection requests
to occupy the instantiable report control blocks of
the IED *.18.50.201 64154 *.18.50.64 102

<1> 2015-12-14 19:30:29 SCADA-61850-IDS IDS MBD-2-a
suspicious destination address of the SMV packet
*:**:cd:66:40:26 **:**:00:00:40:26

<1> 2015-12-14 20:15:45 SCADA-61850-IDS IDS MPD-1
Abnormal switch-in signal *.18.50.16 102 *.18.50.3
42018

<2> 2015-12-14 20:25:23 SCADA-61850-IDS IDS MPD-2
Abnormal differential current of protection relay
*.18.50.18 102 *.18.50.5 45302
```

Fig. 11. Alert examples in the log file

TABLE I
SCADA-Specific IDS Comparisons

IDS	Application scenarios	Protocols	Implementation methods	Implementation tool	Process time	Accuracy
[9]	Power plants	Modbus TCP	Critical state analysis	C#	< 1 ms	99%
[12]	IEC 61850 substations	ARP/ICMP/HTTP/FTP/Telnet	Blacklist rules	Snort	Not published	100%
[14]	Digital substations	MMS/GOOSE	behavior-based detection	sensor equipment	Not published	98.89%
[15]	Digital substations	GOOSE/SMV	Anomaly detection	C/C++	< 0.5 ms	99.81%
Proposed SCADA-IDS	IEC 61850 substations	IEC 61850 (MMS/GOOSE/SMV)	Multidimensional IDS (ACD+PWD+MBD+MPD)	ITACA (C/C++)	< 0.3 ms	100%

VI. CONCLUSION

Compared with physical security for conventional substations, and cybersecurity for IT networks, research on

intrusion detection for IEC 61850 based substations is lacking. In particular, published literature lacks validation of solutions using data from real electrical substations. Furthermore, and as

a result, many published approaches do not focus on providing solutions that are truly tailored to practical implementation at the physical application layer. This research has proposed and developed a multi-layered IDS that focuses on the specific physical environment and application data of the substation to be protected. Key to this is the novel use of configuration information from the SCD file in order to automatically configure the deployed IDS to the substation where the IDS is installed. The proposed solution also adopts detection approaches based around expert knowledge such as GOOSE and SMV parameter configuration data. This provides a clear advantage over existing proposals that are more generic in nature and do not take account of the practical operational environment. The proposed IDS has been implemented and validated in a realistic substation environment. The proposed IDS offers a significant advancement in protecting modern substations against the growing threat of targeted cyberattacks against electrical infrastructure. The IDS has been deployed in a real 500kV smart substation as a trial application. Future work will focus on gathering useful operation data and obtaining practical experience, for further refinement of the system.

VII. REFERENCES

- [1] *Communication Networks and Systems in Substations*, IEC Std. 61850, 2003.
- [2] Y. Yang, K. McLaughlin, S. Sezer, T. Littler, et al., "Multiattribute SCADA-Specific Intrusion Detection System for Power Networks," *IEEE Trans. on Power Delivery*, vol. 29, pp. 1092-1102, Jun. 2014.
- [3] D. Kushner, "The real story of Stuxnet," *IEEE Spectrum*, vol. 50, pp. 48-53, Mar. 2013.
- [4] M. J. Assante, "Confirmation of a Coordinated Attack on the Ukrainian Power Grid." SANS Institute, Bethesda, USA, Jan. 2016. [Online]. Available: <http://ics.sans.org/>
- [5] *Power Systems Management and Associated Information Exchange – Data and Communications Security*. IEC Std. 62351.
- [6] J. Hoyos, M. Dehus, and T. X. Brown, "Exploiting the GOOSE protocol: A practical attack on cyber-infrastructure," in *Proc. 2012 IEEE Globecom Workshops*, pp. 1508-1513.
- [7] C. W. Ten, J. Hong, and C. C. Liu, "Anomaly detection for cybersecurity of the substations," *IEEE Trans. Smart Grid*, vol. 7, pp. 865-873, Dec. 2011.
- [8] S. Cheung, B. Dutertre, M. Fong, U. Lindqvist, K. Skinner, and A. Valdes, "Using model-based intrusion detection for SCADA networks," in *Proc. SCADA Security Scientif. Symp.*, 2007, pp. 127-134.
- [9] A. Carcano, A. Coletta, M. Guglielmi, M. Masera, I. N. Fovino, and A. Trombetta, "A multidimensional critical state analysis for detecting intrusions in SCADA systems," *IEEE Trans. Ind. Inf.*, vol. 7, pp. 179-186, May. 2011.
- [10] I. N. Fovino, A. Carcano, T. De Lacheze Murel, A. Trombetta, and M. Masera, "Modbus/DNP3 state-based intrusion detection system," in *Proc. 24th IEEE Int'l Conf. on Adv. Inf. Netw. and Appl.*, 2010, pp. 729-736.
- [11] R. R. R. Barbosa, R. Sadre, and A. Pras, "Flow whitelisting in SCADA networks," *Int'l Journal of Critical Infrastructure Protection*, vol. 6, pp.150-158, Aug. 2013.
- [12] U. K. Premaratne, J. Samarabandu, and T. S. Sidhu, "An intrusion detection system for IEC61850 automated substations," *IEEE Trans. Power Del.*, vol. 25, pp. 2376-2383, Oct. 2010.
- [13] U. Adhikari, T.H. Morris, P. Shengyi, "A cyber-physical power system test bed for intrusion detection systems," in *Proc. 2014 IEEE PES General Meeting*, pp.1-5.
- [14] K. YooJin, K. Huy Kang, L. Yong Hun, and L. Jong In, "A behavior-based intrusion detection technique for smart grid infrastructure," in *Proc. 2015 IEEE PowerTech*, pp. 1-6.
- [15] H. Junho, L. Chen-Ching, and M. Govindarasu, "Integrated Anomaly Detection for Cyber Security of the Substations," *IEEE Trans. Smart Grid*, vol. 5, pp. 1643-1653, Jul. 2014.
- [16] Y. Yang, K. McLaughlin, T. Littler, S. Sezer, B. Pranggono, and H. F. Wang, "Intrusion Detection System for IEC 60870-5-104 based SCADA networks," in *Proc. 2013 IEEE Power and Energy Society General Meeting*, pp. 1-5.
- [17] H. Junho, L. Chen-Ching, and M. Govindarasu, "Detection of cyber intrusions using network-based multicast messages for substation automation," in *Proc. 2014 IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, pp. 1-5.
- [18] H. Yoo and T. Shon, "Novel Approach for Detecting Network Anomalies for Substation Automation based on IEC 61850," *Multimedia Tools and Applications*, vol. 74, pp. 303-318, Mar. 2014.
- [19] N. Moreira, E. Molina, J. Lázaro, E. Jacob, and A. Astarloa, "Cyber-security in substation automation systems," *Renewable and Sustainable Energy Reviews*, vol. 54, pp. 1552-1562, Nov. 2015.
- [20] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber-Physical System Security for the Electric Power Grid," *IEEE Proc.*, vol. 100, pp. 210-224, Jan. 2012.
- [21] C.-C. Liu, A. Stefanov, J. Hong, and P. Panciatici, "Intruders in the Grid," *IEEE Power Energy Magazine*, vol. 10, pp. 58-66, Jan. 2012.
- [22] H. Hadeli, R. Schierholz, M. Braendle, and C. Tuduca, "Generating configuration for missing traffic detector and security measures in industrial control systems based on the system description files," in *Proc. 2009 IEEE Conf. on Technologies for Homeland Security*, pp. 503-510.
- [23] Y. Yang, H. T. Jiang, K. McLaughlin, L. Gao, Y.B. Yuan, W. Huang, and S. Sezer, "Cybersecurity Test-Bed for IEC 61850 based Smart Substations," in *Proc. 2015 IEEE Power and Energy Society General Meeting*, pp. 1-5.
- [24] *Communication networks and systems for power utility automation—Part 8-1: Specific communication service mapping (SCSM) – Mappings to MMS (ISO 9506-1 and ISO 9506-2) and to ISO/IEC 8802-3*, IEC Std. 61850, 2011.
- [25] *Communication networks and systems for power utility automation—Part 9-2: Specific communication service mapping (SCSM) – Sampled values over ISO/IEC 8802-3*, IEC Std. 61850, 2011.
- [26] *Communication networks and systems in substations—Part 6: Configuration description language for communication in electrical substations related to IEDs*, IEC Std. 61850, 2004.
- [27] Y. Ren, Z. Wang, X. Tang, L. Wang, Y. Du, et al., "Testing system for substation automation system based on IEC61850," in *Proc. 2011 Int'l Conf. on Advanced Power System Automation and Protection*, pp. 2396-2399.
- [28] RFC 2030, Simple Network Time Protocol (SNTP) Version 4, IETF, [Online]. Available: <http://www.ietf.org>.
- [29] J. Hurley, A. Munoz, and S. Sezer, "ITACA: Flexible, scalable network analysis," in *Proc. 2012 IEEE Int. Conf. on Commun. Ind. Forum & Exhibit.*, pp.1084-1088.
- [30] D. Kuipers, & M. Fabro, "Control systems cyber security: Defense in depth strategies". United States. Department of Energy. 2006.
- [31] W. L. Wang M. H. Liu, X. C. Zhao and G. YANG, "Shared-network scheme of SMV and GOOSE in smart substation," *Journal of Modern Power Systems and Clean Energy*, vol. 2, pp. 438-443, 2014.